DOCUMENT RESUME

ED 062 891 FL 003 235

AUTHOR Bordie, John G.
TITLE Language Tests and Linguistically Different Learners: The Sad State of the Art.
PUB DATE [72]
NOTE 37p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Diagnostic Tests; Educational Strategies; Language Proficiency; *Language Research; *Language Skills; *Language Tests; Linguistic Competence; Linguistic Performance; *Measurement Instruments; *Norms; Placement; Predictive Ability (Testing); Standardized Tests; Standard Spoken Usage; Teaching Techniques; Test Construction; Test Reliability

ABSTRACT

Present school programs place considerable reliance on the results of standardized language tests both as placement and diagnostic devices, yet the validity of these tests is open to question. Certain questions arise in light of current theory and research. To what extent are currently available measuring techniques useful for identifying the characteristics of linguistically different learners? Are they helpful in planning instructional strategies? How might the learning potentialities of linguistically different learners be measured? What are the high-priority test needs? Two main branches of investigation exist: the individual's ability in language as contrasted with his ability in standard language. Current understanding of language, language learning, and curriculum design indicates some confusion of goals in these three areas which must be clarified before test findings may be used with the same meaning in each area. (Author/VM)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

John G. Bordie
Associate Professor of Linguistics
and Curriculum & Instruction and
Associate Director, Foreign
Language Education Center
The University of Texas at Austin

ED 062 891

Language Tests and Linguistically Different Learners: The Sad State of the Art

Research

The number of programs to aid students who are linguistically different from the "average" student has had an extensive and rapid growth in recent years. These programs hope to supply specialized training for such students which will allow them to continue in school at approximately the same pace as the majority of the school population. It has been hoped that by doing so the student will be on an equal footing with his fellows at the termination of his school years and will be able to compete for his place in society with no unusual handicaps. In some instances, students who are representative of the national norm are in the minority and the school program must make provision for their abilities without benefitting these students at the expense of the others.

With the availability of various instructional techniques and methods, our requirements are for tests and other measuring instruments which will allow identification of areas requiring support and subsequent proper placement in appropriate classroom or school groupings. Many techniques in daily use could be more effectively exploited if sufficient accurate information were available to teachers and curriculum planning specialists. In addition to availability, the information should be in a form which may be easily interpreted by all likely users rather than remaining solely in the domain of the test specialist.

IL 003 235

Many courses of action suggest themselves. It is possible to establish a single pattern of instruction to which all students must adapt; but too often where such a plan has been established the result is that students adapt to the curriculum without benefitting directly from the subject area content. Other programs suggest highly individualized, highly specialized instruction designed to meet the needs of the individual student. Unfortunately, as evidenced in current practice, this latter course is as wasteful of time and money as the first is of natural talent.

Complicating the problem is the factor of trait visibility. Some student traits are very apparent, e.g., the student who is deaf or blind or who has emotional problems, which we have learned through long experience to handle with some success. Other traits are not so apparent, e.g., the student with a different native language or cultural background or a dialect or socio-economic class level which is different from the school norm. We have not yet learned how to handle these differences with any degree of satisfaction for the benefit of the student or society.

These less visible traits are the ones to which great importance must be attached for they underlie all surface efforts in education. Is this pupil ahead of his peer group because he is naturally intelligent or because his family is wealthy enough to provide him with all of the academic advantages? Is that pupil behind his peer group because he is naturally dull or because he cannot understand the language used in the classroom? Is one student favored because

he sounds like the teacher and another student ignored because he uses a different variety of speech? Many tests have been devised in an attempt to answer these questions.

To what extent are these tests useful and what characteristics do they identify of the various linguistically different groups? In the absence of a generally accepted standard of language use to which instructional techniques may relate or which can serve as an effective model, most tests and measuring techniques have been handicapped. How can one measure variance from the norm or those characteristics which define the norm when no two persons agree on what the definition should be? Any standard which has been established has not usually been sufficiently widespread to be useful.

The definition recently stated by an interdisciplinary working committee (22:4) would appear to sidestep many criticisms regarding standard language definitions. It provides a relational model in the language frequently used on network radio-television newscasts. Yet many characteristics of English as spoken by Negroes, Mexican-Americans, Puerto Ricans, Louisiana French-Acadians, and other minority language groups are generally lacking in the group characteristics used as a reference. One has only to recall the speeches of President Kennedy and President Johnson to recognize the limitations of many definitions.

Consider the anecdote of the elderly rancher in Texas who was pleased with President Johnson because "at last he had a President who spoke standard English" in the light of research by McDavid (31)

which indicated that both whites and Negroes interpreted Southern speech as being substandard English whether spoken by a middle or low socio-economic scale speaker. This same research indicated that stress, intonation, and pitch along with all associated para-language gestures were more indicative of language ability than any of the other usual language characteristics normally thought significant for measurement, such as syntax, vocabulary, grammar, and so on. Additionally, when Negroes were asked to identify the race of an unseen speaker using either standard or non-standard language, the users of non-standard were identified as Negro (3).

How useful are the various techniques and what do they measure? To a considerable degree most predictive measuring techniques conform to an uncertainty principle similar to that used in statistics and physics: one can predict the behavior of a group with fair reliability, but the techniques used for predicting group performance with considerable success may be relatively useless for the prediction of individual behavior, for the individual is not bound by the same sum-total of constraints which limit the group.

Prediction of individual behavior is much like the problem of buying a ready-made suit: if you are average, you will be well satisfied; if you are not average, you will have a strange or an awkward fit. A standard of measurement with 90 percent reliability would appear to be very satisfactory for predictive purposes, but the characteristics of the group remaining, usually from the extremes of the range, are very unclear. It is these groups at the limits of predictive reliability which are of considerable interest at this time.

Although it has been demonstrated that grade level and age are highly correlated, r=.98 (21), the remaining percentage of the school population is of sufficient numerical size in the total population to warrant special treatment. They are most commonly those students who are linguistically or culturally different.

There is little question that tests and other measuring techniques are available to measure a wide variety of skills and abilities. One can demonstrate that tests are reliable instruments which measure very accurately and consistently. Yet though these tests are available and the scores which they supply have been standardized for use in a wide variety of situations, one must ask the question along with Page "not whether it measures, for what it measures is very well measured, but whether it is measuring the right thing" (37).

Intelligence tests have recently been withdrawn from use in the school systems of New York City, Washington, D. C., and Los Angeles, California, because intelligence as measured by current tests is a particular grouping of abilities that has been singled out from the total range of mental abilities required for satisfactory performance in our society. It is entirely possible that the range of abilities which has been singled out is irrelevant to an accurate map of intelligence. Also, to a considerable extent most tests of IQ are loaded with factual material generally known only by certain groups and the scores on the test may be more indicative of teacher or examiner expectations than of individual examinee ability. Recent research indicates that the IQ score (as currently secured) is not

predictive of learning ability in students low on the socio-economic scale although the IQ score may be highly predictive for students from the middle ranges of the socio-economic scale (40).

Some varieties of early age tests appear to test different abilities from those abilities measured by tests for older age groups (6:B). Head Start programs appear to be more effective with children with low IQ scores than those with middle IQ scores; this despite the fact that the type of skills emphasized in Head Start programs are supposed to be those skills which the mid-range children master and excel in during later school years.

It is possible that maturation rates are different for the various groups and that low IQ scores or low socio-economic levels are predictive of a faster maturation rate which levels off much sooner than does the slower maturation rate of the middle level child (16;19). The exact implications of this are unclear at the moment for it may be that one variety of test measures ability to learn while another variety measures the amount of actual knowledge.

Most tests measure very many of the same things: language mechanics, e.g., punctuation, spelling, capitalization; formal grammar; recognition of correct form; vocabulary; reading comprehension; usage; parts of speech; sentence types; editorial revision; and recognition of differences in style. Except for oral reading sections in some test batteries, therefore, the tests are printed tests. They can only measure written English and are incapable of measuring oral English.

Most tests also require a "correct" response in answer to a question and assume that the "correct" response is a reflection of

standard English. All tests of this sort discriminate against speakers of a second language, speakers of regional dialects, and less verbal lower achievers. To the extent that they confuse written language with oral language or assume that only written language ability is indicative of language ability in general, they will be misleading and useless. Where tests assume that editorial ability is indicative of writing ability, they are presumptuous and limited devices. The ability to handle language cannot be equated with the ability to handle a sample of the language since ability may be partial in parts of the language, complete in some areas, and totally lacking in very many others. A fluent conversationalist is not necessarily a good orator and is even less likely to be a good novelist.

The ability to verbalize responses to questions about language ability is more indicative of the mastery of a formal pattern of language analysis (e.g., knowledge of traditional grammar, ability to comprehend the test, and knowledge of formal descriptive terms or the degree to which the student's public language corresponds to the formal standard language used in the school. Regardless of geographic area, low socio-economic level students have a public language which they use at home or at play less like that of the formal language used in the school and in print than is the language of middle socio-economic level students (30).

When the low socio-economic student is asked to identify an item as incorrect, it is extremely likely the "incorrect" form is the form he uses in his public speech or which is common in his dialect. When

he is asked to restructure a sentence or a paragraph, his lack of contact with a variety of styles puts him at a distinct disadvantage to the middle socio-economic level student who has read more widely or has traveled a bit throughout the United States (4). If the student speaks a different language, his performance in English will logically be much improved if he is given initial instruction in his own language prior to undertaking a test in the second language. Obviously, if he cannot understand what he is to do, he cannot perform satisfactorily (34). Additionally, when the examiner is of the same ethnic background as the student, the understanding and confidence which is created allows the student to perform far better than if he were being tested by someone of different background (34).

Language abilities which can be measured consist of a number of widely varying skills. Since some of the skills are indicative of the ability to handle the production of material while others indicate the ability to handle the reception of material, it may be useful to arrange these varieties of skills in a convenient matrix illustrating both production and reception.

Chart I

Phase Matrix of Language Skills*

Production Encoding		Reception Decoding	
Speaking	Writing	Listening	Reading
Semantics	Semantics	Semantics	Semantics
Syntax	Syntax	Syntax	Syntax
Morphology	Morphology	Morphology	Morphology
Lexicon	Lexicon	Lexicon	Lexicon
Phonemes	Graphemes	Phonemes	Graphemes
Kinesics	Paralanguage	Kinesics	Paralanguage

*Adapted from Cervenka (15) and MacNamara (29)

Clearly, preschool children and illiterates of all ages will be completely lacking, or nearly so, in the skills listed under writing and reading. Bilingual students will have a dual-matrix situation in which the relation of capacity in one area of the native language matrix to the same area of the second language matrix must be considered. This may be a very difficult, perhaps impossible, task. "No easy way of measuring or characterizing the total impact of one language on another in the speech of bilinguals has been, or probably can be, devised" (48).

When such abilities are translated from the terminology of the linguist, a variety of characteristics emerge which include items ranging from the completely mechanical skills to the most abstract cognitive abilities. It is of considerable importance that by far the most common skills tested are the purely mechanical despite the lack of evidence of any correlation between such skills and language proficiency. From a sample of two hundred educational experiments regarding abilities which were considered significant as predictors, the following emerged when items with a citation incidence of five or more were tabulated. The purely mechanical skills were excluded since they were mentioned in virtually every report and no correlation could be found for them.

Varied and Flexible Vocabulary
Aural Comprehension
Oral Usage
Phonetic Accuracy
Length and Number of "T" Units (main clause with associated subordinate clause)
Cloze Ability (ability to comprehend or reconstruct material from which every nth element is removed)
Frequency of Use of Tentativeness, Relational Words, Conditional Clauses, and Optional Grammatical Patterns
Ability to Restructure or Rephrase
Ability to Handle Syntactic Cues

Associated research with specific interest in these areas is detailed (1; 4;5;26;27;35;46). These items depend to a considerable extent on standardization and norming procedures. Few tests have extensive norms and those that do have such norms may be completely irrelevant to curriculum needs. There is some research which gives no indication of a person's socialization or intellectual identity from the observable presence of a linguistic characteristic such as the use of standard verb forms, preferred lexical items, or acceptable pronunciation (45).

Tests

More tests which are in current use in research are tests which have been developed for the particular area to be investigated or the specific experiment in which they will be used. Most researchers are apparently dissatisfied with published tests and are convinced that they must develop their own which are more suited to their needs. The above sample of research (chosen from studies reported in Research in Education under the ERIC descriptors of comparative testing, language testing, language capacity, and disadvantaged students) referred to twenty-nine published tests, the most frequently cited meriting only five citations.

All other tests cited were specifically developed tests produced for the research in which they were used. Some of these tests are now commercially available and will be going through the norming and use process necessary for validation. It should be emphasized that the following list is not comprehensive for many of the tests were

unavailable at the time of writing. Explicit and complete reviews for many of the following tests may be found in the Mental Measurements Yearbooks (6-8) and Tests in Print (9). References in parentheses at the end of each summary refer to the more complete reference in the appropriate source material.

A final point to be noted: since curricula change and students change, the older a test, the more satisfactory a local group will appear when measured against national norms; obviously if one wants to demonstrate that a new methodology has solved a particular problem, one should use an older test rather than the newest test to appear. Such new tests tend to measure areas not previously considered in the methodology while the older tests measure those items which the methodologies have had sufficient experience and practice in solving and teaching.

1. Barrett-Ryan-Schrammel English Test: two citations. Designed to survey student proficiency in English mechanics, facilitate grouping and placement, and to diagnose deficiencies.

Items are based on the common content of leading textbooks and courses of study; which and how many are not specified. No evidence is available to indicate that the test performs better than a reasonably good teacher-made examination and no evidence is provided that placement is facilitated by test scores to a greater extent than previous grades in English (7:B).

2. California Achievement Test: two citations.

Designed to survey student achievement in reading, arithmetic, and language.

Test items are well constructed although their coverage is somewhat limited. Population sample size is rather small and achievement scores are based on material produced in 1957 (8:A).

3. California Language Tests: one citation.

Designed to test mechanics, word usage, sentence structure, and verbal expression.

Standardization is poorly defined and the reviewer is not convinced of its representativeness, particularly since school system testing for the South and the Northwest is completely lacking. The word usage section appears to be particularly prone to dialect misinterpretation (8:G).

4. Cooperative English Test: one citation.

Designed to test usage, grammar, mechanics, sentence structure, spelling, and vocabulary.

The test does not examine the ability to write or speak and resembles tests found in English workbooks. The test was produced in 1938 and has been only slightly changed since that time (8:H; 8:I).

5. Cooperative School and College Ability Test: one citation. Designed to test the total range of academic achievement for placement in high school and college.

The test places a heavy emphasis on reading and writing. It predicts academic achievement better for women than for men, and the total

score predicts English grades far better than the partial sub-score on verbal ability. The test is not wide enough in range and tends to discriminate against lower achievers (8:N).

6. Differential Aptitude Test: one citation. Designed to test general verbal aptitude on basis of responses indicating agreement or disagreement with test item.

Test usage has not changed since 1947. The norms appear large, but are actually quite small when each level is considered. The ends of the normal range are apparently unbalanced and, possibly, biased (8:S).

7. Gloria and David Beginning English, Series No. 20, Test 6, Language Arts-- Spanish-English: one citation. Designed to test phonology, oral comprehension, and aural comprehension. To a lesser degree, also tests translation ability.

Local norms must be developed through use of check sheet supplied. Designed to be used in conjunction with filmstrips and tape recorder in order to supply same testing situation to all students (47).

8. Essentials of English Test: one citation. Tests usage, spelling, mechanics, sentence structure, and vocabulary.

The test is essentially based on material developed for use in 1939 and is relatively out-of-date by current standards. Vocabulary appears to be overly formal (8:J).

9. Frostig Development Test of Visual Perception: two citations. A clinical tool for the definition of perceptual abilities and levels of understanding.

The test relies on verbal responses to a series of pictures which are somewhat culturally biased. Primarily used with handicapped children, such as aphasics and emotionally disturbed, and lower age level children. Has a lesser degree of validity with non-handicapped children. Directions to the child reflect some language and culture problems which are not considered in the scoring (8:R).

10. Greene-Stapp Language Abilities Test: two citations. Test items focus on the determination of abilities related to classroom content, such as spelling, grammar, and mechanics.

Student must register his understanding of the correctness of an item or its converse by assigning a grammatical reason for the error. This ability as well as the ability to recognize formal vocabulary may discriminate against the low socio-economic level or nonstandard dialect student (7:C).

11. Illinois Test of Psycholinguistic Ability: five citations. Designed to differentiate facets of cognitive ability by means of 346 items classed into twelve groups: auditory reception; visual reception; visual sequential memory; auditory association; auditory sequential memory; visual association; visual closure; verbal expression; grammatic closure; manual expression; auditory closure; and sound blending.

This is the most commonly used test in the research explored. Only a few studies are available of the recent revision, but a significant

number of projects are underway including the use of the test in the evaluation of national Head Start programs. It should be noted that the size of various subsections of the test varies from the maximum of fifty items for auditory reception to four for verbal expression. The average number of items per facet is twenty-nine. Validation of the test was done on groups of white children in Wisconsin. Whether the test is satisfactory is yet to be determined, but on the surface the test does not inspire confidence. Most teachers indicate that the lack of ability in verbal expression is the most serious disadvantage their students can have in the classroom. The ITPA diagnoses such ability on the basis of four items from a total of 346. How such a diagnosis can be accurate or predictive except in the most general of terms is a mystery (8:Q).

12. Iowa Test of Basic Skills: one citation.

13. Iowa Test of Educational Development: two citations. Both tests investigate language skills through vocabulary, reading comprehension, spelling, mechanics, and usage.

The testee is expected to find the error in the material presented to him in these tests. As a result, the tests emphasize editorial abilities and are more indicative of experiential background than of skills or language ability. Dialect variation would appear to be a factor in low scoring. The standard sample reflects this for there is a pronounced shortage of trials in urban Southern schools (8:B).

14. Metropolitan Achievement Test: four citations.

15. Metropolitan Readiness Test: four citations.

Both tests are designed to test achievement and ability in the total range of skills thought appropriate for success in the classroom.

The tests are conservative tests reflecting the understanding the publisher had of the national curriculum at the time the tests were constructed. Norms were based on white children in twenty-six states.

Language material is inadequate for the student at the lower end of the scales since chance scores are less than one-half grade below the minimum level for which norms are offered. The use-ranges introduce a reliability of 6/9 or 66 percent which is not sufficient to discriminate with sufficient confidence at the low end of the scale. Since the publisher specifically warns against the use of the tests for individual diagnosis, it is difficult to understand why this test is as popular as it is unless one believes the test users assume their school population to be uniformly middle class and white (6:A;8:C;8:T).

16. Objective Test in Grammar: two citations.

Measures recognition and verbal ability in terms of traditional formal grammar.

The publisher provides no data on reliability, no data on validity, no manual, no norms, no standards, and the test is untimed. Presumably the test users have used the test for some time and are comfortable with it on the basis of local experience (6:A;9).

17. Oral Directions Test: one citation.

Designed to test aural comprehension, visual discrimination, and motor skills of individuals over sixteen years of age.

The test finds its greatest usefulness as a screening device for low level factory positions. The test discriminates against rural populations and, apparently, women (7:D).

18. Peabody Picture Vocabulary Test: three citations.

Designed for use with children incapable of taking regular tests, e.g., cerebral palsied, brain damaged.

The examiner reads a word and the subject indicates a picture. The test is short. Standards supplied with the test are based on children living near Nashville, Tennessee. The score point jumps are very large: 50 points indicates IQ score of 101 at age 5.5 but a score of 89 at age 5.6 (8:0).

19. Purdue High School English Test: one citation.

Tests vocabulary, spelling, reading comprehension, and mechanics.

The subject matter coverage is restricted and there are few indications of curricular relevance. A peculiarity of the test is that chance scores put the individual into high percentile ranks on subtests (45th percentile, 64th percentile) (6:A,8:K).

20. Science Research Associates High School Placement: one citation.

21. Language Arts Tests: two citations. Both tests are designed to evaluate a student's ability in

total curriculum areas and in the specific language skills requisite for academic success.

The tests attempt to evaluate total student ability as well as language proficiency, but it is possible to achieve a grade equivalent increase of one-half to one year and fifteen to twenty percentile points solely for one correct spelling. The validity of the tests are open to question since several possible alternate answers may be given for many items on the test without being considered in the total score. "Only if the ablest students take the tests in the spirit of 'what answer did the author intend to be considered right' instead of 'what answer or answers can be defended as correct alternatives' would they score well on the test" (8:D:88;8:L).

22. Sequential Test of Educational Progress: two citations. Designed to test the entire range of educational ability.

The test provides for the typical student only and discriminates against students at both ends of the ability range. The test is an attractive one and provides scores which give the appearance of being informative. Most reviewers are uncertain of the worth of the information provided, however (7:A;8:E).

23. Stanford Achievement Test: three citations. A test for the basic range of school achievement (reading, spelling, arithmetic, language, social studies, and science).

The language portion of the test asks students to identify words which are the best descriptors of a picture, e.g., to choose items

of similar phonic value. In so doing, the test is concerned with the measurement of ability in formal school standard written English. Such ability is usually lacking among the linguistically different. Apparently the continued wide use of the test comes from the user's long familiarity with it and the development of valid local standards. Since the tests are based on the texts and the curricula of the 1950's, they will measure what was modern then and is, perhaps, standard now (8:F).

24. Stanford-Binet IQ Scale: four citations.

With the publication of Arthur R. Jensen's article "How Much Can We Boost IQ and Scholastic Achievement" / (23), most IQ tests and scales were immediately labeled as suspect, biased, racist, misleading, and worse. Such tests are characterized by the investigation of the ability to initiate and maintain abstract reasoning as well as the ability to solve problems of various natures. Jensen claims that this ability or group of abilities is inherited. It is difficult to evaluate the pros and cons of the argument from the point of view of language but three items are worth noting: (1) learning ability is characterized by at least a two-level system consisting of associative or rote learning and conceptual or cognitive learning; (2) most curricula stress conceptual learning to the almost complete exclusion of associative learning; and (3) language teaching has emphasized the value of associative-rote learning in the early stages of language acquisition. It would follow that IQ tests as currently conceived would discriminate against the individual still in an associative learning stage. Whether this is

true or not, it should be further noted that the Stanford-Binet norms were established in 1937 and are now some thirty-three years old (8:P;18;23).

25. English Usage Test for Non-Native Speakers of English: two citations.

26. Michigan Test of English Language Proficiency: two citations.

27. Listening Test for Students of English as a Second Language: two citations.

28. Test of English as a Foreign Language: two citations.

29. Oral Rating Form for Rating Language Proficiency in Speaking and Understanding English: two citations. These five tests are designed to evaluate pronunciation, grammar and word order, vocabulary, aural comprehension, sentence length, and mechanics. In addition, some also evaluate general speed of speech and oral fluency.

These tests are standardized on the basis of a student's ability to function at the collegiate level. It is assumed that all persons taking the test are more than sixteen years of age and are intent on an academic program at or beyond the freshman level in college. The language which is rated is highly specific to general academic usage. While these tests function reasonably well, they are not suitable as ability scales, being more precisely accurate for placement purposes. Here the validity varies considerably and is highly dependent on standardization within the local situation. One could take the tests

off the shelf and use them immediately, but their predictive ability would have to be taken on faith (8:M).

Implications

Are these tests and their findings helpful in planning instructional strategy? Since any test will help the teacher or the curriculum developer by providing a score or an evaluation which functions as a reference point, we can say that they are helpful in planning instructional strategy. But the strategy which is planned may be considerably different from existing strategy and may require wholesale change in the curriculum to be completely effective. Again, since most test developers caution against the use of subsections of a test as diagnostic tools, few tests can be said to be satisfactory.

A number of tests which do make provision for such diagnostic use provide very limited and quite unnatural samples of the range of abilities normally expected of the student. The result of such partial diagnostic work may conceivably hinder the student or misdirect his energies. Such tests and their diagnoses are valid only when the user has had sufficient experience with them in the situation in which they are to be used so that he feels generally confident with their predictions and can interpret the test scores with ease. Once this occurs he can make satisfactory diagnostic use of them. Then the cycle starts once more: is that which is being tested relevant to what is being taught; and can a partial sample of the student's work predict his ultimate level of achievement? The answer to date appears to be that one should wear the shoe if it fits properly.

Some results have been determined by several of the tests which indicate the validity of one or another instructional strategy. Many linguists and language researchers have argued that the difference between the linguistically sophisticated and the linguistically immature is not so much the awareness of correct and incorrect usage but rather the general knowledge of a wide range of language varieties and adequate contact with the varieties most characteristic of school instruction. Many experiments and tests indicate that it is far more fruitful to expand the student's language repertory than it is to "correct" the language he uses in his daily life. By providing a wide range of experiential contacts, the teacher and the curriculum can make clear that language consists of a variety of styles which must be mastered, each of which has its own value and use. Such awareness is highly effective in providing the individual with a solid base for his later cognitive understanding of use levels (14;32;33;38;39).

How the repertory is to be expanded is another and more difficult matter. Some specialists have argued that language growth in the early stages is an associative process requiring a fairly large amount of rote learning. Language teaching specialists have made effective use of this effect in their development of the audio-lingual methodology which stresses the development of habitual patterns of behavior. When research has examined this area and has investigated the language acquisition of children, there has been considerable evidence that rote or associative learning is quite important in the early stages of the language learning process (11;19;23;24;25;43;47).

It is possible that the benefits derived from such teaching are due mainly to the structuring of the content which is introduced as a consequence of the development of appropriate materials. Many of the reports indicate that structuring is quite important (11;24; 25;36) and advocate the teaching of topics from the point of view of the final examination. It is suggested that the final examination be a comprehensive map of all those skills and abilities which the student is expected to have at the end of his period of study (13).

Some types of instruction and instructional techniques have been indicated as having limited or marginal value in the language teaching process. While they are suitable in average circumstances, these techniques are apparently of little use in the exceptional circumstances found in the linguistically different classroom. These are the use of diagramming (17), instruction in standard English as an aid to reading for the linguistically different (41), instruction in formal language at the early stages of language learning (30), and instruction centered on the school situation alone (39).

Some techniques are successful. From the numerous reports of such successes, one might think that basic discoveries in the teaching process are being made. In reality, these positive reports only echo a rather obvious characteristic of education: when the classroom situation is familiar to the student, the curriculum content apparent, the teacher enthusiastic, and student needs and interests met, then a wonderful amount of learning is achieved.

Most positive reports indicated that a few specific techniques are of more than passing value: (1) earlier instruction for the linguistically different is likely to prevent many language problems in later school years (12); (2) pre-instruction in the student's language or dialect helps him perform better in standard English (34); (3) the use of content materials to teach language is more successful than the use of language materials alone (44); (4) highly structured programs are more successful with the linguistically different than they are with the standard population (24;25); and (5) the earlier the language material is introduced to the student, the more likely it is that he will master it (10;28;38). However, although many individuals argue that a specific age is more advantageous for the introduction of instruction in language, there is no evidence that one age is better than any other. All research has indicated is that an early start gives more practice, and practice is apparently what is necessary for language mastery (10).

What are the high priority test needs? A few items are apparent as possible topics: (1) we must be able to measure an individual's competence in language (whether he speaks a nonstandard dialect or another language) as contrasted with his competence in standard English; (2) we need a convenient checksheet so that teachers and school administrators can determine what standard of language is used in school or is used in the community; (3) we must have an acceptable definition of standard English which allows for the richness of some of the dialects spoken in the United States; (4) we need tests which

distinguish between language proficiency and degree of socialization; (5) we need to know what it is that is required for satisfactory performance in the school curriculum in language other than that performance solely based on written language; and (6) finally, but not least of all, we need a definition of language which takes into account all of the abilities used in human communication. We are still an inordinate distance from a satisfactory definition, let alone a detailed specification of skills.

The work being done in the Nevada desert with the chimpanzee Washoe (20) indicates that language is not solely confined to man and may be used for significant interspecies communication. If the Gardners are successful with Washoe, we will be in dire need of a true definition of "linguistically different."

Summary

Present school programs place considerable reliance on the results of standardized language tests both as placement devices and as diagnostic devices, yet the validity of these tests is open to question. This paper discusses the following four questions in the light of current theory and research:

1. To what extent are currently available measuring techniques useful for identifying the characteristics of linguistically different learners;
2. Are they helpful in planning instructional strategies;

3. How might the learning potentialities of linguistically different learners be measured; and

4. What are the high priority test needs?

General findings of research indicate the general preference of most researchers for the specially developed test designed for the particular research at hand. More than half of all programs surveyed used such specially developed tests.

The results of such testing indicate two main branches of investigation as fruitful areas: the individual's ability in language as contrasted with his ability in standard language. Current understanding of language, language learning, and curriculum design indicates some confusion of goals in these three areas which must be clarified before test findings may be used with the same meaning in each area.

Publication Sources for Tests

1. Barrett-Ryan-Schrammel English Test

   Harcourt, Brace and World, Inc., 757 Third Avenue,
   New York, New York 10017

2. California Achievement Test

   California Test Bureau, DelMonte Research Park,
   Monterey, California 93940

3. California Language Tests

   California Test Bureau, DelMonte Research Park,
   Monterey, California 93940

4. Cooperative English Test: Usage, Spelling, and Vocabulary

   Cooperative Test Division, Educational Testing Service,
   Princeton, New Jersey 08540

5. Cooperative School and College Ability Test

   Cooperative Test Division, Educational Testing Service,
   Princeton, New Jersey 08540

6. Differential Aptitude Test

   Psychological Corporation, 304 E. 45th Street,
   New York, New York 10017

7. Gloria and David Beginning English, Series #20
   Test 6 - Language Arts - Spanish-English

   Language Arts, Incorporated, 1205 W. 34th Street,
   Austin, Texas 78705

8. Essentials of English Test

   American Guidance Service, Inc., 720 Washington Avenue, S. E.,
   Minneapolis, Minnesota 55414

9. Marianne Frostig Developmental Test of Visual Perception

   Consulting Psychologists Press, 577 College Avenue,

   Palo Alto, California 94306

10. Greene-Stapp Language Abilities Test

   Harcourt, Brace and World, Inc., 757 Third Avenue,

   New York, New York 10017

11. Illinois Test of Psycholinguistic Ability

   University of Illinois Press, Urbana, Illinois 61803

12. Iowa Test of Basic Skills

   Houghton Mifflin Co., 2 Park Street,

   Boston, Massachusetts 02107

13. The Iowa Tests of Educational Development

   Science Research Associates, 259 E. Erie Street,

   Chicago, Illinois 60611

14. Metropolitan Achievement Test

   Harcourt, Brace and World, Inc., 757 Third Avenue,

   New York, New York 10017

15. Metropolitan Readiness Test

   Harcourt, Brace and World, Inc., 757 Third Avenue,

   New York, New York 10017

16. Objective Test in English (Grammar)

   Perfection Form Company, 214 W. Eighth Street,

   Logan, Iowa 51546

17. Personnel Tests for Industry: Oral Directions Test

   Psychological Corporation, 304 E. 45th Street,

   New York, New York 10017

18. Peabody Picture Vocabulary Test

American Guidance Services, Inc., 720 Washington Avenue, S. E.,
Minneapolis, Minnesota 55414

19. The Purdue High School English Test

Houghton Mifflin Company, 2 Park Street,
Boston, Massachusetts 02107

20. S. R. A. High School Placement Test

Science Research Associates, 259 E. Erie Street,
Chicago, Illinois 60611

21. S. R. A. Achievement Series: Language Arts Tests

Science Research Associates, 259 E. Erie Street,
Chicago, Illinois 60611

22. Sequential Test of Educational Progress

Cooperative Test Division, Educational Testing Services,
Princeton, New Jersey 08541

23. Stanford Achievement Test

Harcourt, Brace and World, Inc., 757 Third Avenue,
New York, New York 10017

24. Stanford-Binet IQ Scale

Houghton Mifflin and Company, 2 Park Street,
Boston, Massachusetts 02107

25. English Usage Test for Non-Native Speakers of English

26. Listening Test for Students of English as a Second Language

27. Oral Rating Form for Rating Language Proficiency in Speaking and Understanding English

American Language Institute, 3065 O Street, N. W.,
Washington, D. C. 20007

28. Michigan Test of English Language Proficiency

Follet's Michigan Bookstore, 322 S. State Street,

Ann Arbor, Michigan 48104

29. Test of English as a Foreign Language

Educational Testing Services, Princeton, New Jersey 08540

Bibliography

1. Andrade, Manuel and others, Measurement of Speaking Skills in Elementary Level Spanish Instruction - Denver-Stanford Project on the Context of Instructional Television. Denver Board of Education: N. D. E. A. Report 7A-354, 1963. ERIC ED 018153.
2. Baratz, Joan C., Language and Cognitive Assessment of Negro Children: Assumptions and Research Needs. Paper read at American Psychological Association Convention, 1968. ERIC ED 022157.
3. Baratz, Joan C., "A Bidialectal Test for Determining Language Proficiency." ERIC Research in Education, 1968. ED 020519.
4. Barritt, Loren S. and others, A Comparison of the Psycholinguistic Functioning of "Educationally Deprived and Educationally Advantaged" Children. Ann Arbor: University of Michigan Center for Research on Language and Language Behavior, 1965. ED 022537.
5. Bennett, Stanley and others, The CLOZE Procedure - A Review of the Literature and Possible Application to the Study of Deviant Language Functions. Ann Arbor: University of Michigan Center for Research on Language and Language Behavior, 1967. ED 017920.
6. Buros, O. K., The Fourth Mental Measurements Handbook. Highland Park, New Jersey: The Gryphon Press, 1953, A. p. 311; B. p. 48.
7. Buros, O. K., The Fifth Mental Measurements Handbook. Highland Park, New Jersey: Gryphon Press, 1959, A. pp. 62-75; B. pp. 330-32; C. pp. 344-47; D. pp. 501-02.

8. Buros, O. K., The Sixth Mental Measurements Handbook. Highland Park, New Jersey: The Gryphon Press, 1964, A. pp. 17-21; B. pp. 48-55; C. pp. 55-67; D. pp. 88-93; E. pp. 100-09; F. pp. 109-28; G. pp. 544-46; H. pp. 550-51; I. pp. 551-56; J. pp. 565-67; K. pp. 574-77; L. pp. 577-79; M. pp. 638-41; N. pp. 714-18; O. pp. 820-23; P. pp. 828-32; Q. pp. 852-53; R. pp. 854-57; S. pp. 1001-07; T. pp. 1073-74.

9. Buros, O. K., Tests in Print. Highland Park, New Jersey: The Gryphon Press, 1961, p. 67.

10. Carroll, John B. "Foreign Languages for Children-What Research Says," The National Elementary Principal, 39 (May, 1960).

11. Carroll, John B., "The Contributions of Psychological Theory and Educational Research to the Teaching of Foreign Languages," Modern Language Journal, 49 (May, 1965),pp. 273-81.

12. Carter, John L., The Long Range Effects of a Language Stimulation Program upon Negro Educationally Disadvantaged First Grade Children. Houston, Texas: University of Houston, 1967. ED 013276.

13. Cartier, Francis A., "Criterion-Referenced Testing of Language Skills," TESOL Quarterly, 2 (March, 1968).

14. Cazden, Courtney B., Some Implications of Research on Language Development for Preschool Education. U. S. O. E. Report BR50215-29, 1966. ED 011329.

15. Cervenka, Edward J., "The Measurement of Bilingualism and Bicultural Socialization of the Child in the School Setting: The Development of Instruments," in Head-Start Evaluation and Research. Austin, Texas: Child Development Evaluation and Research Center, The University of Texas, 1967, pp. 10-11.

16. Daniel, Artie A. and Douglas E. Giles, A Comparison of the Oral Language Development of Head-Start Pupils with Non Head-Start Pupils. U. S. O. E. Report BR 68293, 1966. ED 010848.

17. Desousa, Albert M. and Milly Cowles, "An Experimental Study to Determine the Efficacy of Specific Training in Listening." Paper read at the February Convention of the American Educational Research Association, 1967. ED 016656.

18. Edson, Lee, "'JENSENISM,'n. The Theory that IQ is Largely Determined by the Genes," The New York Times Magazine, (August 31, 1969) pp. 10-11, 40-41, 43-47.

19. Entwisle, Doris R., Developmental Socio-Linguistics. Baltimore, Maryland: Johns Hopkins University, U. S. O. E. Report BR1610-1, 1967. ED 011611.

20. Gardner, R. A. and B. T. Gardner, "Teaching Sign Language to a Chimpanzee," Science, 165 #3894 (August 15, 1969) pp. 664-72.

21. Hill, Edwin C. and Margaret K. Hill, "Written Language Development of Intermediate Grade Children." Pittsburgh: University of Pittsburgh, 1966, U. S. O. E. Report CRP1571. ED 010059.

22. Horn, Thomas D. (ed.), Reading for the Disadvantaged: Problems of Linguistically Different Learners. Sponsored by the International Reading Association. New York: Harcourt, Brace and World, Inc., 1970.

23. Jensen, Arthur R., "How Much Can We Boost IQ and Scholastic Achievement," Harvard Educational Review, 39 (Winter, 1969) pp. 1-123.

24. Karnes, Merle B. and others, A Comparative Study of Two Pre-School Programs for Culturally Disadvantaged Children - A Highly Structured and a Traditional Program. Urbana: University of Illinois, U. S. O. E. Report BR5-1181, 1966. ED 016524.

25. Kellogg, Ralph E., A Study of the Effect of a First-Grade Listening Instructional Program upon Achievement in Listening and Reading. San Diego County: California School Board, U. S. O. E. Report BR68469, 1966. ED 012232.

26. Klima, Ursula Bellugi, Evaluating the Child's Language Competence. Urbana: University of Illinois National Laboratory on Early Childhood Education, 1968. ED 019141.

27. Loban, Walter, Language Ability - Grades Ten, Eleven, and Twelve. Berkeley: University of California, 1967. ED 014477.

28. Macleish, Andrew and William R. Seat, Materials and Methods for Teaching Structural and Generative Grammar to High School Students and Their Teachers. DeKalb: Northern Illinois University, 1967, U. S. O. E. Report BR51112. ED 015173.

29. MacNamara, John, "The Bilingual's Linguistic Performance - A Psychological Overview," The Journal of Social Issues, 23 (April, 1967) pp. 58-59.

30. Massad, Carolyn E., A Comparative Study of Language Aptitude and Intelligence in Sixth-Grade Children From Low Socio-Economic and Middle Socio-Economic Levels. Paper read at the Convention of the American Education Research Association, 1968. ED 020291.

31. McDavid, Raven I. and William M. Austin, Communication Barriers to the Culturally Deprived. Chicago: University of Chicago, 1966, U. S. O. E. Report CRP2107. ED 010052.

32. McNeil, John D. and James C. Coleman, Auditory Discrimination Training in the Development of Word Analysis Skills. Los Angeles: University of California at Los Angeles, 1967. ED 018344.

33. Melaragno, Ralph J. and Gerald Newmark, A Pilot Study to Apply Evaluation Procedures in First Grade Mexican-American Classrooms. Santa Monica: California Systems Development Corporation, 1968, Report TM3930-000-00. ED 026170.

34. Mycue, Elena I., Testing in Spanish and the Subsequent Measurement of English Fluency. M. A. Thesis. Denton, Texas: Texas Women's University, 1968. ED 026193.

35. O'Donnell, Roy C., "An Objective Measure of Structural Complexity in Children's Writing." Paper read to the American Educational Research Association, 1968. ED 016534.

36. Osborn, Lynn R., "Speech Communication and the American Indian High School Student," The Speech Teacher, 17 (January, 1968) pp. 38-43. ED 018426.

37. Page, F. B., A review of the Iowa Test of Educational Development in Buros' The Sixth Mental Measurements Handbook. (op. cit.) p. 51.

38. Politzer, Robert, An Investigation of the Order of Presentation of Foreign Language Grammar Drills in Relation to their Explanation. Stanford, California: Stanford University, 1967. ED 018163.

39. Prindiville, Sister Francis De Sales, "Pupil-Teacher Relationships: Major Factors in Developing Language Arts in Slow Learners." Champaign, Illinois: National Council of Teachers of English, 1967. ED 017515.

40. Rohwer, Jr., W. D., Social Class Differences in the Role of Linguistic Structures in Paired Associate Learning, Elaboration and Learning Proficiency. Berkeley: University of California, 1967. ED 016246.

41. Rystrom, Richard, The Effects of Standard Dialect Training on Negro First Graders Learning to Read. ~~Diablo Valley~~ Concord, California: Diablo Valley College, 1968. ED 029717.

42. Savard, Jean-Gay, Bibliographie Analytique de Tests de Langue (In French with English key). Quebec: Laval University, 1969.

43. Skinner, B. F. Verbal Behavior. New York: Appleton-Century-Crofts, Inc., 1957.

44. Stemmler, Anne O., The Psychological and Cognitive Aspects of Teaching English as a Second Language. San Antonio, Texas: U. S. O. E. Report BR50249-18, 1966. ED 011606.

45. Stolz, Walter S., The Role of Dialect in the School Socialization of Lower Class Children--Head Start Evaluation. Austin: The University of Texas, 1967. ED 019121.

46. Taft, Jerome and Melvin Tennis, The Development of a Test to Assess the Occurrence of Selected Features of Non-Standard English in the Speech of Disadvantaged Primary Children. Miami: Dade County Public Schools, 1968. ED 015790.

47. Taylor, Thomasine, A Comparative Study of the Effects of Oral-Aural Language Training on Gains in English Language for Fourth and Fifth Grade Disadvantaged Mexican-American Children. Ph.D. Dissertation. Austin: The University of Texas at Austin, 1969, pp. 65-68.

48. Weinreich, Uriel, Languages in Contact. New York: Linguistic Circle of New York, 1954, pp. 63-64.